Agriculture, Environment & Society

Journal homepage: www.aes.uoz.ac.ir

Detection of sugar content in sugar beets using hyperspectral imaging

Mohsen Mahdiani [a], Mehdi Khojastehpour [*b], Mahmood Reza Golzarian [c]

[a] Ph.D. Student, Department of Biosystems Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
[b] Department of Biosystems Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
[c] School of Engineering and IT, Murdoch University, WA 6150, Australia

ARTICLE INFO

Article history:
Received: 4 October 2024
Accepted: 15 November 2024
Available online: 1 June 2025

Keywords:
Hyperspectral imaging
Rapid detection
Regression analysis
Sugar beet
Sugar content

ABSTRACT

Measuring the sugar content in sugar beet is challenging because of the labor, expenses, and chemicals required. Hence, there is a necessity for a more efficient method to measure this content. This study aims to establish an efficient, non-invasive method for detecting sugar content in sugar beets using hyperspectral imaging, which could revolutionize quality control in the sugar beet industry. This study used hyperspectral imaging to analyze 400-950 nm sugar beet paste. Pre-processing techniques such as SNV (Standard Normal Variate) and SG (Savitzky-Golay), along with wavelength selection methods such as SPA (Successive Projection Algorithm) and CARS (competitive adaptive reweighted sampling), were applied. Furthermore, various regression models including MLR (multiple linear regression), PLS (Partial Least Squares regression) , and SVR (Support Vector Regression) were employed for prediction. Evaluating these models based on R^2 and RMSE criteria, the PLS regression model with SNV pre-processing and SPA wavelength selection stood out, achieving an R^2 value of 0.91% and RMSE of 0.24. These findings suggest the potential of hyperspectral imaging as a rapid and accurate means of determining sugar content in sugar beet across the VIS-NIR spectrum.

Highlights

- The analysis of hyperspectral images was investigated to predict the sugar content.
- SG and SNV algorithms were used for preprocessing.
- CARS and SPA algorithms were used to select the effective wavelengths.
- Models for prediction have been created by MLR, PLS and SVR algorithms.
- The best model for prediction of sugar content were selected.

1. Introduction

Assessing the sugar content in sugar beet is essential for both sugar factories and farmers, particularly in regions where payments to farmers are based on sugar content (Uygan et al., 2021; Wang et al., 2022). Since sucrose comprises C–H, O–H, C–C, and C–O bonds, various methods such as enzymatic analysis, HPLC, and polarimetry have been employed for their measurement (Pan et al., 2015a). Among these, the polarimetric method is favored for its simplicity and is commonly used in sugar factories. This method involves mixing 26 grams of sugar beet paste with 177 grams of lead acetate solution to measure sugar content through polarization (Roggo et al., 2004; Shabani and Sepaskhah, 2019).

Researchers have concentrated on developing precise and rapid methods for determining sugar content without relying on chemical reagents. Spectroscopy techniques, including hyperspectral imaging, have emerged as effective solutions. One of the pioneering studies in this area is by Roggo (Roggo et al., 2004), who employed an NIR spectrometer covering the 400 to 2498 nm wavelength range. This study demonstrated that spectroscopy can serve as a fast and straightforward method for assessing the sugar content in sugar beets. Additionally, the study highlighted the potential of this technique for simultaneously identifying other parameters. In Pan's research (Pan et al., 2015a; Pan et al., 2013; Pan et al., 2015b), two portable spectrometers with working ranges of 400-1100 nm and

* Corresponding author.
E-mail address: mkhpour@um.ac.ir
https://doi.org/10.22034/jelsa.2024.473159.1081

900-1600 nm were employed to obtain spectral data from the beets in intact and sliced conditions. The study, which considered interactance mode, consistently confirmed that sugar content could be accurately identified under all tested conditions (sliced beet and intact beet). In a separate study conducted by Babaei (Babaee et al., 2019), a near-infrared (NIR) spectrometer was employed in the wavelength range of 942-1576 nm in interactance mode to investigate the effects of spectroscopic location and sugar beet skin on the spectrometer's performance. Utilizing ANOVA analysis, the results indicated that within this spectral range, the wavelength of 1393 nm exhibited the highest correlation. Furthermore, spectral measurements taken from four locations—Northwest (NW), Southwest (SW), Northeast (NE), and Southeast (SE)—approximately one centimeter above and below the root neck, where the root diameter is greatest, did not yield significant differences. However, a notable distinction was observed between the spectroscopy of the skin and flesh of the sugar beet root, with measurements from the flesh demonstrating greater accuracy. Also in Bagharpur's study (Bagherpour et al., 2014), the quantification of sugar content was validated through the application of infrared spectroscopy within the wavelength range of 900-1600 nm on sugar beet samples, along with the implementation of pre-processing techniques.

Hyperspectral imaging is a technique that enables the extraction of internal compositional information from materials through imaging. This approach has demonstrated promising outcomes in numerous studies focused on agricultural products aimed at identifying their internal constituents. For instance, research conducted on mulberry fruits (305-1090nm wavelength in reflectance mode) has indicated that this method can effectively quantify anthocyanin levels (Li et al., 2023), while another study assessed the soluble solids content in Tribute Citru using the same technique (400–1000 nm wavelength in reflectance mode) (Li et al., 2023). In investigations carried out by Pan (Pan et al., 2014; Pan et al., 2016a), hyperspectral imaging was utilized on sugar beet slices within the spectral range of 500-1000 nm, resulting in relative reflectance data that demonstrated considerable accuracy. These results suggest that hyperspectral imaging may be an effective substitute for traditional methods of determining sugar content. In Wang's study (Wang et al., 2022), he effectively employed an Unmanned Aerial Vehicle (UAV) to evaluate the sugar levels in fields by combining hyperspectral imagery with RGB, multispectral, and thermal infrared images. His findings demonstrated that hyperspectral imaging can be used by UAVs for remote sensing to accurately measure sugar content. Nonetheless, many factories are unable to adopt these technologies widely due to their high costs thus posing problems of accessibility hence limited usage. As such, the cost-effective ways of applying hyperspectral imaging in determining sugar content levels of beet paste could greatly impact sugar processing companies. This would lead to more accurate measurements in the sugar factories.

In light of the research context, several pertinent questions arise: Firstly, is it feasible to ascertain sugar content through hyperspectral imaging of sugar beet paste? Secondly, can a regression-based methodology be developed for accurately determining sugar levels in sugar beet? How can we enhance identification accuracy by implementing efficient wavelength selection algorithms? The present study aims to address these inquiries.

2. Material and methods

2.1. Sampling and imaging

An automated machine randomly samples a portion of the sugar beets delivered to the factories. These beets undergo a process of washing, shredding, and combining to create a uniform mixture. This mixture is then analyzed to determine the sucrose content and other elements (Fasahat et al., 2022). In the Torbat-Heydarieh Sugar factory located in Torbat-Heydarieh, Khorasan Razavi, Iran, a total of 150 samples were prepared using this method. Each sample was divided into two segments: one segment was sent to an imaging lab (Automation and Computer Vision Laboratory -ACVL) at Ferdowsi University of Mashhad for imaging, while the other segment was sent to the Agricultural Research, Education and Extension Organization (AREEO) in Mashhad, Khorasan Razavi, Iran, to obtain data on sucrose content using a Betalizer (Figure 1), the sugar content (SC) was determined using the polarimetric method.

The linear scan-type imaging device utilized in this study is a Desktop Scanner manufactured by Parto Afzar Sanat Co. (Zanjan, Iran). It possesses several key specifications, including a scan length of 200mm, a scan resolution of 0.05mm, a vertical arm height of 400mm, and four halogen light sources that ensure uniform illumination. The imaging spectral range of this device spans from 400nm to 950nm, with a spectral interval of 2nm between consecutive spectral images. To ensure accurate calibration, the device was calibrated by imaging a white standard plate, which serves as the device standard, and obtaining a dark reference image by covering the lens with the cap. In addition to its ability to manually store and display the average spectrum of a selected region, the device also can store spectroscopic images. The primary focus of this study is the analysis of the stored images. The images within the device create a spectroscopic cube with dimensions of 114 x 36 x 516. Here, 516 represents the spectral resolution of the image, which corresponds to the number of acquired spectra ranging from 400nm to 950nm. Furthermore, the values 36 and 114 represent the spatial resolution—specifically the width and height—of the captured images. For data analysis, MATLAB software (VER 2022b, MathWorks Inc, US) was employed on an ACER ASPIRE laptop equipped with 20GB RAM and a core i7 CPU processor.

After preparation, each sample was placed in a tightly sealed container with minimal air exchange and stored in a refrigerator at -4 degrees Celsius. The samples were then quickly transported to the laboratory, taking less than 24 hours. A precise 2 cm stainless-steel cube mold was created to ensure standardization in laboratory analysis. This mold was used to shape the mixed sugar beet paste into uniform

sample blocks, ensuring all samples had identical dimensions and height to maintain consistency during sampling. The samples were carefully inserted into the molds and flattened with a knife to address any potential height variations during the sampling process (Figure 2). Without this mold, smoothing the surface would have required significant effort, and the color would have changed in the process. However, the small dimensions of the mold allowed for quick smoothing with a knife. If the sample's color changed during sampling, some material was removed from the middle and placed into the mold. Determining the appropriate sample size was essential to achieve a sample with a specific thickness quickly. Additionally, applying pressure at the end of the mold and removing some material from the sample's surface allowed for faster imaging and minimized color changes. Each sample was sampled at least twice, and the average of the spectra was used as the final sample spectrum. Subsequently, the sample was positioned beneath the imaging device, and a region of interest (ROI) measuring 51x51 pixels was designated at the center of each sample surface using the associated imaging software application. The spectral images across all spectral bands were then captured from this defined ROI.

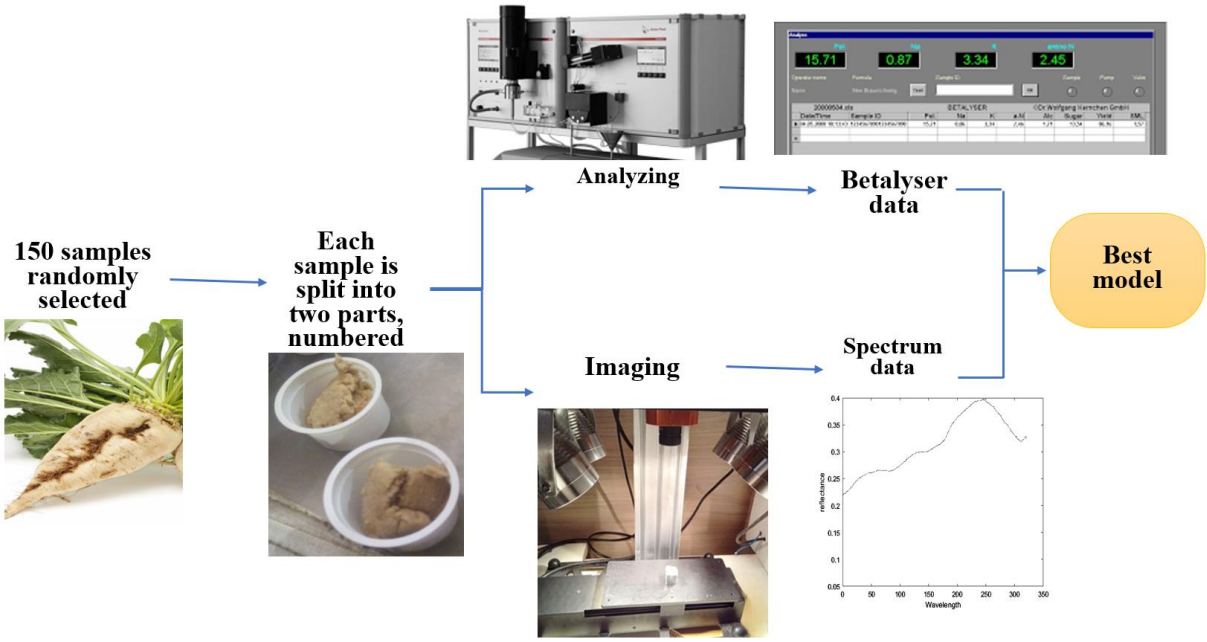


Figure 1. Flowchart of model building

Figure 2. A 2x2x2 cm sampling mold and a manual rectangular Region of Interest (ROI) on the sample after the spectral image is captured

2.2 Data processing

In spectroscopic devices, the existence of device noise in both low and high spectral ranges necessitates the exclusion of specific spectra during the processing stage. For instance, in a particular device, the initial and final spectra were eliminated as reported by Yang et al. (Yang et al., 2014), leading to a total of 321 spectra out of the original 516 spectra captured. The remaining 321 spectra were then utilized in the subsequent processing procedures, as illustrated in Figure 3. In subsequent analyses, we utilized data cubes, each with dimensions of 321 × 51 × 51 and comprising a total of 150 samples—a complete dataset.

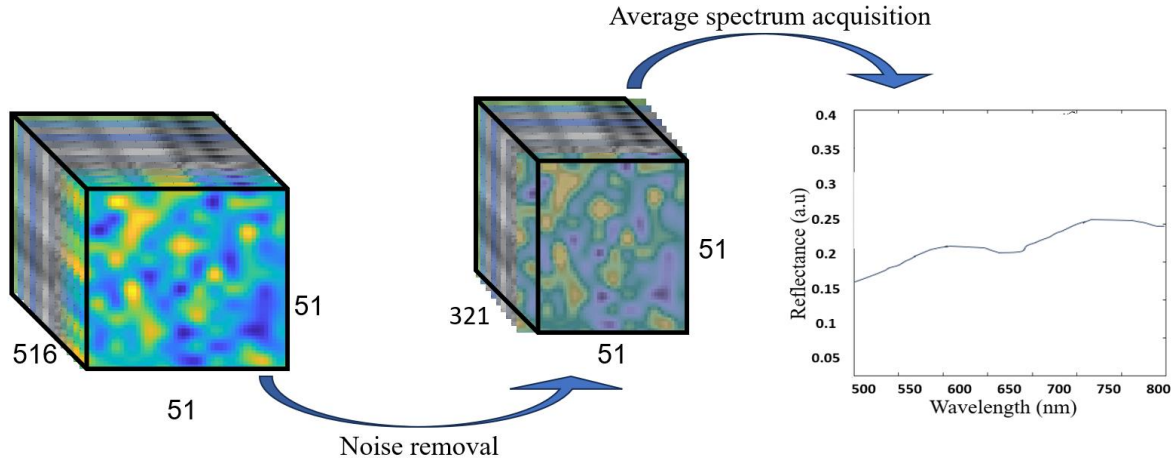


Figure 3. Noise removal and average spectrum acquisition

To address the impact of outliers on predictive models caused by instrument and operational errors, the researchers in this study employed the Monte Carlo Partial Least Squares (MCPLS) method (Zhang et al., 2015). To begin with, a calibration set was created by randomly selecting sample fractions, while the remaining fractions were designated as the prediction set. PLS models were then generated repeatedly to ensure that each sample was utilized multiple times in the prediction set. As a result, each sample produced a series of predictive residual errors (PRE). For each sample in the prediction set, the Mean of Predicted Residual Errors (MPRE) and the Standard Deviation of Predicted Residual Errors (STDPRE) were calculated (Figure 4) (Guo et al., 2012).

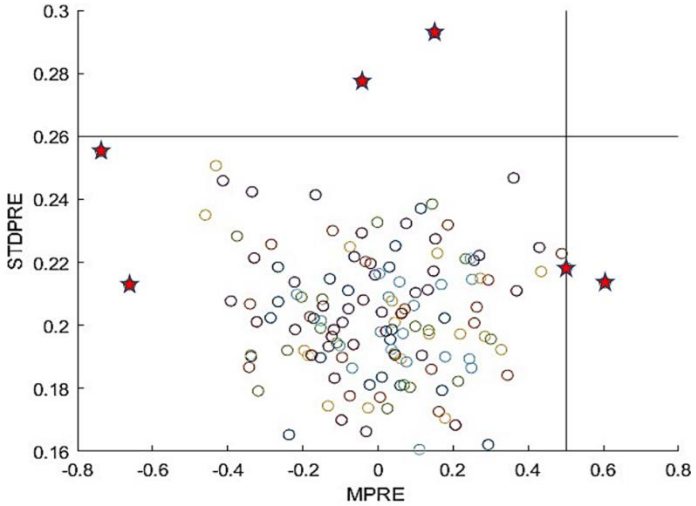


Figure 4. MPRE versus STDPRE plot for determining outliers

2.2.1. Preprocessing

Data acquired from spectroscopic imaging systems can be influenced by various factors, including light scattering resulting from the interaction between the sample and the detector, changes in sample size, surface roughness of the sample, noise generated due to an increase in system temperature, and other factors. These unintended factors have the potential to affect the accuracy of models and lead to significant discrepancies in the interpreted data. These elements introduce additional noise that may mask true spectral signals, thus compromising the reliability of analytical outcomes (Zhang et al., 2021). Ensuring the integrity of the data not only strengthens calibration models but also fosters greater confidence in subsequent interpretations and applications across various scientific fields. Therefore, to achieve stable, accurate, and reliable validation models, it is necessary to preprocess the data (Rossel, 2008). In this particular study, two common

preprocessing techniques (Sim et al., 2024), namely Standard Normal Variate (SNV) and Savitzky-Golay (SG) smoothing, were utilized. SNV and SG are. These preprocessing techniques are employed to mitigate the influence of multiplicative effects such as baseline shifts and intensity changes in spectral data. SNV involves transforming each data point by subtracting the mean and dividing it by the standard deviation across all data points. On the other hand, Savitzky-Golay (SG) smoothing is used to remove random noise while preserving important features of the captured spectra. The SG smoothing technique applies a convolution process with a set of predefined coefficients to a moving window of data points. SNV was performed with a degree of 3 and a window size of 5. The SG method, which utilizes a window size of 5 and a degree of 3, is widely employed for noise reduction (Yu et al., 2021). SNV, on the other hand, is useful in eliminating the scattering effect caused by light and particle size. Furthermore, the impact of these preprocessing methods on the performance of the predictive model was evaluated and compared with the performance metrics obtained when only the raw data, without any preprocessing, was fed into the model.

2.2.2. Selection of effective wavelengths

To enhance modeling accuracy and reduce computational time, employing different algorithms for selecting effective wavelengths is crucial. This study investigated two commonly used methods: Competitive Adaptive Reweighted Sampling (CARS) and Successive Projection Algorithm (SPA). The CARS method, inspired by the concept of "survival of the fittest" from Darwinian Theory, utilizes Monte Carlo sampling to select a subset of wavelengths through iterative processes and competition. This method evaluates wavelengths based on their absolute coefficient value in PLS modeling by optimizing variables with an exponential decreasing penalty function and adaptive weighted resampling (Ji et al., 2020). The subset with the smallest RMSEV is identified as characteristic wavelengths, obtained through repetitions of sampling (Shao et al., 2020).

On the other hand, the SPA is a valuable tool for forward variable selection, enhancing modeling speed and accuracy by reducing collinearity and extraneous information among variables. SPA operates by selecting the optimal wavelength with the highest design value in each iteration until the desired number of wavelengths is reached. The number of variables is determined by minimizing the root mean square error (RMSE) in multiple linear regression (MLR) calibration (Shao et al., 2020), contributing to improved modeling outcomes.

In general, the SPA method involves adding one wavelength at a time based on correlation, while the CARS algorithm samples and re-weights wavelengths according to their performance in predicting the target property. This study aims to compare these two methods in terms of their effectiveness and efficiency in spectral data analysis.

MLR is a widely used technique for establishing a linear relationship between spectral data and chemical components. Despite being an older algorithm, MLR remains efficient, especially when the number of features is less than the observations (Fei et al., 2023). Additionally, PLSR is a method that utilizes full-spectrum data to predict sample composition by identifying the highest covariance between reference values and spectral data. PLSR combines various statistical analysis methods to model the relationship between dependent and independent variables, making it particularly useful in scenarios with more variables than observations and when multicollinearity is present among the X values (Nie et al., 2023). Support Vector Regression (SVR) is another supervised Machine Learning technique that focuses on structural risk minimization and demonstrates high-performance (Ahmad et al., 2020), especially with multidimensional data and high-dimensional datasets (Mesut et al., 2023).

Given the wide utilization and satisfactory outcomes of these three regression methods in hyperspectral image studies (Mitku et al., 2024; Shao et al., 2020; Soltanikazemi et al., 2022), all three methods were employed in this study to assess the most effective prediction models (Figure 5).

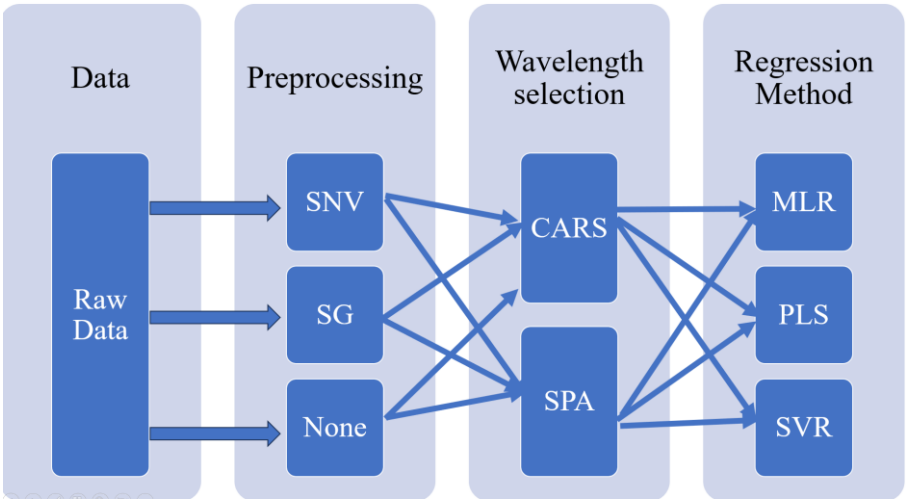


Figure 5. Workflow of finding regression method

2.3. Model evaluation

In this study, due to the limited amount of data, the cross-validation method was employed. The data was split into two groups: 80% for training and 20% for validation. Regression model evaluation metrics include the correlation coefficient (R) and root mean square error (RMSE). R indicates the goodness of fit, while RMSE measures the difference between predicted and actual values (eq. 1-2). A value of R^2 close to 1 suggests high stability and fit, whereas an RMSE close to 0 indicates stronger predictive ability (Li et al., 2023):

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - y_m)^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (2)$$

In these equations,

y_i = measured value of the attribute for the ith sample in the category;

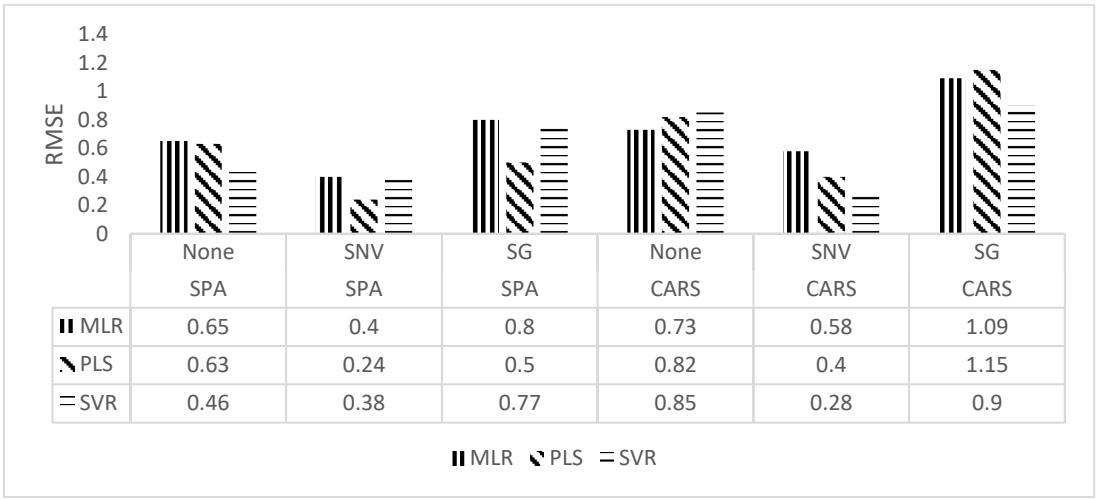
$\hat{y}_i$ =predicted value of the attribute for the ith sample in the category;

y_m= The average measured values of attributes in the category;

n= the number of category samples;

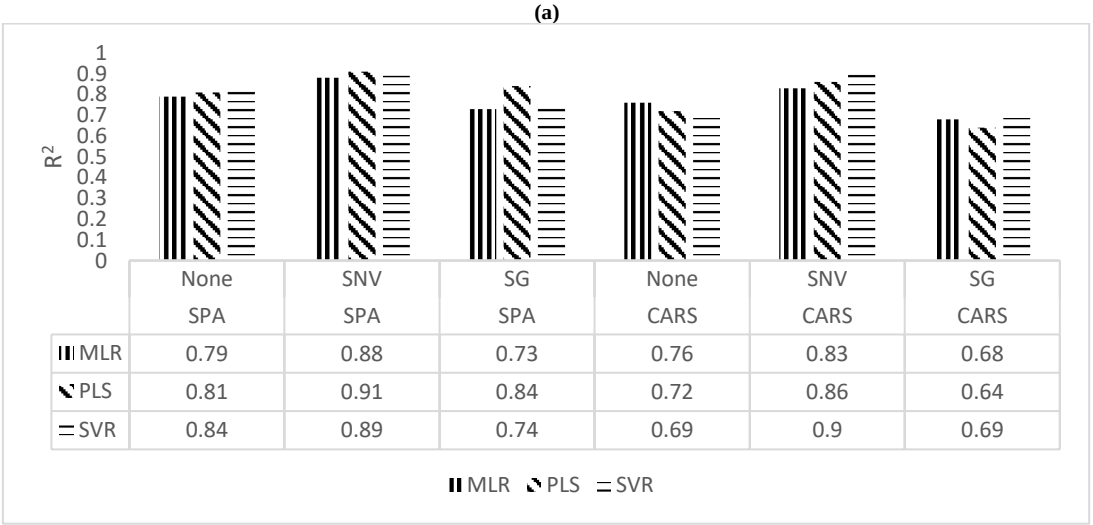
3. Results

After analyzing the Betalyser result shows that the SC ranges from 10.10 to 23.65%. This range is consistent with the typical SC range of 10 to 24 % in sugar factories, indicating a close resemblance to real-world samples. In the spectrum (Figure 3), there is a noticeable dip at 650 nanometers and a peak at 620 nm, associated with the chlorophyll level in the paste (Governici et al., 2017; Tian et al., 2019). This range is also recognized by Pan et al. as a crucial factor in determining chlorophyll in sugar beet (Tian et al., 2019). Additionally, a noticeable peak at 750 nm, suggested by Pan, is likely linked to the third overtone of water(Pan et al., 2016b) or the fourth overtone of C–H (Wei et al., 2017). The O–H and C–H functional groups are associated with the concentration of specific internal compositions, such as soluble solid contents.



	None SPA	SNV SPA	SG SPA	None CARS	SNV CARS	SG CARS
MLR	0.65	0.4	0.8	0.73	0.58	1.09
PLS	0.63	0.24	0.5	0.82	0.4	1.15
SVR	0.46	0.38	0.77	0.85	0.28	0.9

(a)



	None SPA	SNV SPA	SG SPA	None CARS	SNV CARS	SG CARS
MLR	0.79	0.88	0.73	0.76	0.83	0.68
PLS	0.81	0.91	0.84	0.72	0.86	0.64
SVR	0.84	0.89	0.74	0.69	0.9	0.69

(b)

Figure 6. Comparison of different regression, a) comparison of RMSE b) comparison of R^2

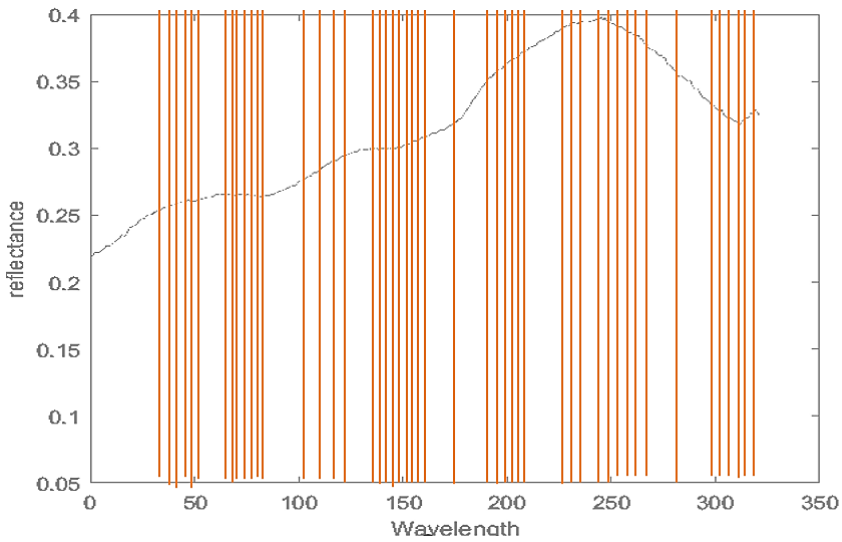


Figure 7. Selected wavelengths for determining sucrose

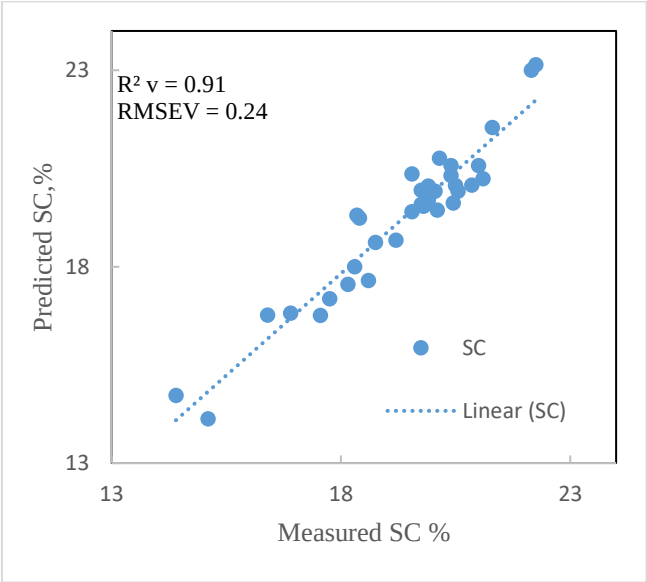


Figure 8. Measured vs predicted value

The MCPLS method was utilized to detect outliers, following a similar approach to Zhang's study (Zhang et al., 2023). For running MCPLS 75% of the selected samples were assigned randomly to the calibration set, whereas the remaining 25% were allocated to the prediction set. This iterative procedure was repeated 5000 times, leading to the generation of a scatter plot that visually depicts the correlation between MPRE and STDPRE. Figure 4 Demonstrates that six samples were identified as outliers. Specifically, any values of MPRE out of range (-0.5, 0.5) and values of STDPRE surpassing 0.26 were classified as outlier data.

The spectral data contained excessive information, linearity, and overlapping, which negatively affected the performance of multivariate calibration models. This study utilized SPA and CARS to identify particular wavelengths with minimal linearity and redundancy, enhancing modeling efficiency.

The number of effective wavelengths was selected after applying effective wavelength selection methods (CARS and SPA) to each type of preprocessing data (None, SNV, SG) then prediction models were created using MLR, PLS, and SVR regression methods, and the results are shown in Figure 6.

In Figure 6, the combination of the SPA and SG algorithms exhibits the highest RMSE value and the lowest R^2 value. Conversely, the combination of the SPA algorithm with SNV demonstrates the lowest RMSE value and the highest R^2 value. Figure 6 illustrates the capability of the SPA algorithm to decrease the RMSE value. This comparison between the SPA algorithm without preprocessing and the CARS algorithm without preprocessing is evident. The best model for identifying effective wavelengths as input for the PLS algorithm is the combination of SNV preprocessing and the SPA algorithm that selects 50 wavelengths (Figure 7). This model achieved a validation R^2 of 0.91, RMSEV of 0.24. Figure 8 also illustrates the performance of the best model for predicting sugar content in the calibration set and validation set.

4. Discussion

In this study, different regression models were employed (PLSR, MLR, and SVR) to analyze hyperspectral data within the visible-near-infrared (vis-NIR) spectral range (400-950nm). The present model successfully determined the sugar content in sugar beet. This study found that the SNV and SG pre-processing algorithms significantly improved model accuracy in most cases, based on validation data. Additionally, algorithms for selecting effective wavelengths, such as CARS and SPA, enhanced model performance. While the CARS algorithm chose fewer wavelengths, its accuracy was lower than SPA. Among the regression methods in this study, MLR and PLS algorithms exhibited better performance than the SVR algorithm in predicting SC. Results are consistent with previous studies (Pan et al., 2014; Pan et al., 2016a) that demonstrated the ability to determine sucrose content within the 500-1000 nm wavelength range using the PLS algorithm with an R^2 of 0.76. Pan's study (Pan et al., 2013; Pan et al., 2015b), revealed that the optimal model for intact beet was 0.81, while for sliced beet it was 0.89, within the wavelength range of 400-1100 nm. One possible reason for the difference in the results could be attributed to the use of sugar beet paste in this study, which provided a more homogeneous sample. Findings align with other studies that determine sugar content in fruit paste and employ hyperspectral images. For instance, Maraphum (Maraphum et al., 2020) identified the amount of sugar in sugarcane within the same wavelength range (400-1000 nm) and showed that using the PLS algorithm with first derivative preprocessing achieved an R2 of 0.66, which is superior to their proposed method. Liu (Liu et al., 2017) demonstrated that using multispectral imaging of tomato paste within the range of 405-970 nm can determine sugar content with an R^2 of 0.93. In this study, the optimal method for predicting the SC level involves employing SNV preprocessing in conjunction with the SPA algorithm to determine the relevant wavelengths for the PLS algorithm. This approach resulted in a validation R2 of 0.91 and a minimal RMSEV of 0.24. These values are significantly higher and lower, respectively than those reported in previous studies (Maraphum et al., 2020; Pan et al., 2016a), indicating the model's higher performance and robustness in predicting sugar content. This research demonstrates, for the first time, that the SC of sugar beet paste can be evaluated through hyperspectral imaging. These findings have promising implications for the agricultural industry, as they can lead to improved crop management practices and enhanced product quality. By leveraging various algorithms such as SPA, PLS, and MLR, scientists can build robust models that provide accurate and reliable results. Furthermore, this study emphasizes the significance of selecting effective wavelengths, with the SPA method emerging as a particularly valuable tool in this regard. Overall, these results contribute to the growing body of knowledge on the application of hyperspectral imaging for quality assessment in agricultural products, highlighting its potential for enhancing efficiency and precision in various industrial processes.

5. Conclusion

The current study sought to measure the sugar content in sugar beet using hyperspectral images of sugar beet paste within the spectral range from 400 to 950 nm. To achieve this, dough samples were taken from the shipments received at the sugar factory. The sugar content of each sample was assessed using a Betalyzer, and hyperspectral imaging of the sugar beet paste was also captured. In this study, pre-processing methods such as SNV and SG were utilized, along with effective algorithms for wavelength selection, specifically CARS and SPA. The results were then fed into MLR, PLS, and SVR algorithms, resulting in the creation of eighteen regression models that combined these approaches. Among the eighteen developed models, the best model was selected with the minimum RMSE and maximum R^2 values. These results indicated that all models were acceptable, but the best performance was achieved when SNV preprocessing was carried out along with the SPA algorithm as an input method for PLS regression. Such results would enable fast, chemical-free evaluation of sugar content, thereby benefiting sugar factories while also assisting farmers in assessing the size of their crops' sugar content. Since this research is quite at the beginning, it would be of great use to extend the project for more general results with more samples involved. Moreover, techniques beyond those applied could still provide further precision. For instance, deep learning techniques may be considered to learn features of the hyperspectral data.

Declaration of Competing Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors also wish to express their gratitude to Torbat_Heydarieh Sugar Factory for their collaboration.

Funding: This work was supported by Ferdowsi University of Mashhad (FUM) [Grant No. 59736].

References

Ahmad, M. S., Adnan, S. M., Zaidi, S., & Bhargava, P. (2020). A novel support vector regression (SVR) model for the prediction of splice strength of the unconfined beam specimens. *Construction and Building Materials, 248*, 118475. https://doi.org/10.1016/j.conbuildmat.2020.118475

Babaee, B., Khanmohammadi, M., Garmarudi, A. B., & Abdollahin Noghabi, M. (2019). Effect of peeling and point of spectral recording on sucrose determination in sugar beet root using near infrared spectroscopy. *Infrared Physics & Technology, 103*, 103065. https://doi.org/10.1016/j.infrared.2019.103065

Bagherpour, H., Minaei, S., & Abdollahian Noghabi, M. (2014). Non-destructive determination of sugar content in root beet by near infrared spectroscopy (nirs). *Journal of Food Science & Technology (2008-8787), 12*(46).

Fasahat, P., Rezaei, J., Sharifi, M., Azizi, H., Fatouhi, K., Mahdikhani, P., Pedram, A., Jalilian, A., & Babaei, B.

(2022). Assessment of root and white sugar yield stability of sugar beet genotypes.

Fei, S., Xu, D., Chen, Z., Xiao, Y., & Ma, Y. (2023). Mlr-based feature splitting regression for estimating plant traits using high-dimensional hyperspectral reflectance data. *Field Crops Research*, *293*, 108855. https://doi.org/10.1016/j.fcr.2023.108855

Governici, J. L., Faria, R. M., dos Reis Tinini, R. C., & Mederos, B. J. T. (2017). Tomatoes maturation analysis with reflectance spectral images. *Journal of Agricultural Science and Technology B.*

Guo, W.-L., Du, Y.-P., Zhou, Y.-C., Yang, S., Lu, J.-H., Zhao, H.-Y., Wang, Y., & Teng, L.-R. (2012). At-line monitoring of key parameters of nisin fermentation by near infrared spectroscopy, chemometric modeling and model improvement. *World Journal of Microbiology and Biotechnology*, *28*, 993-1002. https://doi.org/10.1007/s11274-011-0897-x

Ji, H., Wang, W., Chong, D., & Zhang, B. (2020). Cars algorithm-based detection of wheat moisture content before harvest. *Symmetry*, *12*(1), 115. https://doi.org/10.3390/sym12010115

Li, C., He, M., Cai, Z., Qi, H., Zhang, J., & Zhang, C. (2023). Hyperspectral imaging with machine learning approaches for assessing soluble solids content of tribute citru. *Foods*, *12*(2), 247. https://doi.org/10.3390/foods12020247

Li, S., Song, Q., Liu, Y., Zeng, T., Liu, S., Jie, D., & Wei, X. (2023). Hyperspectral imaging-based detection of soluble solids content of loquat from a small sample. *Postharvest Biology and Technology*, *204*, 112454. https://doi.org/10.1016/j.postharvbio.2023.112454

Li, X., Wei, Z., Peng, F., Liu, J., & Han, G. (2023). Non-destructive prediction and visualization of anthocyanin content in mulberry fruits using hyperspectral imaging. [Original Research]. *Frontiers in Plant Science*, *14*. https://doi.org/10.3389/fpls.2023.1137198

Liu, C., Hao, G., Su, M., Chen, Y., & Zheng, L. (2017). Potential of multispectral imaging combined with chemometric methods for rapid detection of sucrose adulteration in tomato paste. *Journal of Food Engineering*, *215*, 78-83. https://doi.org/10.1016/j.jfoodeng.2017.07.026

Maraphum, K., Saengprachatanarug, K., Wongpichet, S., Phuphaphud, A., & Posom, J. (2020). In-field measurement of starch content of cassava tubers using handheld vis-near infrared spectroscopy implemented for breeding programmes. *Computers and Electronics in Agriculture*, *175*, 105607. https://doi.org/10.1016/j.compag.2020.105607

Mesut, B., Başkor, A., & Aksu, N. B. (2023). Role of artificial intelligence in quality profiling and optimization of drug products *A handbook of artificial intelligence in drug delivery* (pp. 35-54): Elsevier. https://doi.org/10.1016/B978-0-323-89925-3.00003-4

Mitku, A. A., Titshall, L., Zewotir, T., & North, D. (2024). Application of support vector machine regression and partial least-square regression models for predicting sugarcane leaf nutrients content using near infra-red spectroscopy. *Communications in Soil Science and Plant Analysis*, *55*(2), 196-207. https://doi.org/10.1080/00103624.2023.2265426

Nie, B., Du, Y., Du, J., Rao, Y., Zhang, Y., Zheng, X., Ye, N., & Jin, H. (2023). A novel regression method: Partial least distance square regression methodology. *Chemometrics and Intelligent Laboratory Systems*, *237*, 104827. https://doi.org/10.1016/j.chemolab.2023.104827

Pan, L., Lu, R., Tu, K., & Cen, H. (2014). *Detection of moisture, soluble solids, and sucrose content and mechanical properties of sugar beet by hyperspectral scattering imaging.* Paper presented at the 2014 Montreal, Quebec Canada July 13–July 16, 2014.

Pan, L., Lu, R., Zhu, Q., McGrath, J. M., & Tu, K. (2015a). Measurement of moisture, soluble solids, sucrose content and mechanical properties in sugar beet using portable visible and near-infrared spectroscopy. *Postharvest Biology and Technology*, *102*, 42-50. https://doi.org/10.1016/j.postharvbio.2015.02.005

Pan, L., Lu, R., Zhu, Q., Tu, K., & Cen, H. (2016a). Predict compositions and mechanical properties of sugar beet using hyperspectral scattering. *Food and Bioprocess Technology*, *9*, 1177-1186. https://doi.org/10.1007/s11947-016-1710-5

Pan, L., Lu, R., Zhu, Q., Tu, K., & Cen, H. (2016b). Predict compositions and mechanical properties of sugar beet using hyperspectral scattering. *Food and Bioprocess Technology*, *9*(7), 1177-1186. https://doi.org/10.1007/s11947-016-1710-5

Pan, L., Zhu, Q., Lu, R., & McGrath, J. M. (2013). *Detection of sucrose content of sugar beet by visible/near infrared spectroscopy.* Paper presented at the 2013 Kansas City, Missouri, July 21-July 24, 2013.

Pan, L., Zhu, Q., Lu, R., & McGrath, J. M. (2015b). Determination of sucrose content in sugar beet by portable visible and near-infrared spectroscopy. *Food Chemistry*, *167*, 264-271. https://doi.org/10.1016/j.foodchem.2014.06.117

Roggo, Y., Duponchel, L., & Huvenne, J.-P. (2004). Quality evaluation of sugar beet (beta vulgaris) by near-infrared spectroscopy. *Journal of agricultural and food chemistry*, *52*(5), 1055-1061. https://doi.org/10.1021/jf0347214

Rossel, R. A. V. (2008). Parles: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems*, *90*(1), 72-83. https://doi.org/10.1016/j.chemolab.2007.06.006

Shabani, A., & Sepaskhah, A. R. (2019). Optimal amounts of water and nitrogen applied to sugar beet when crop price depends on its sugar content. *Spanish Journal of Agricultural Research*, *17*(3), e1202. https://doi.org/10.5424/sjar/2019173-14487

Shao, Y., Liu, Y., Xuan, G., Wang, Y., Gao, Z., Hu, Z., Han, X., Gao, C., & Wang, K. (2020). Application of hyperspectral imaging for spatial prediction of soluble solid content in sweet potato. *RSC advances*, *10*(55), 33148-33154. https://doi.org/10.1039/c9ra10630h

Sim, J., McGoverin, C., Oey, I., Frew, R., & Kebede, B. (2024). Optimisation of vibrational spectroscopy instruments and pre-processing for classification

problems across various decision parameters. *Food Innovation and Advances*, *3*(1), 52-63. https://doi.org/10.48130/fia-0024-0004

Soltanikazemi, M., Minaei, S., Shafizadeh-Moghadam, H., & Mahdavian, A. (2022). Field-scale estimation of sugarcane leaf nitrogen content using vegetation indices and spectral bands of sentinel-2: Application of random forest and support vector regression. *Computers and Electronics in Agriculture*, *200*, 107130. https://doi.org/10.1016/j.compag.2022.107130

Tian, X., Li, J., Wang, Q., Fan, S., Huang, W., & Zhao, C. (2019). A multi-region combined model for non-destructive prediction of soluble solids content in apple, based on brightness grade segmentation of hyperspectral imaging. *Biosystems Engineering*, *183*, 110-120. https://doi.org/10.1016/j.biosystemseng.2019.04.012

Uygan, D., Cetin, O., Alveroglu, V., & Sofuoglu, A. (2021). Improvement of water saving and economic productivity based on quotation with sugar content of sugar beet using linear move sprinkler irrigation. *Agricultural Water Management*, *255*, 106989. https://doi.org/10.1016/j.agwat.2021.106989

Wang, Q., Che, Y., Shao, K., Zhu, J., Wang, R., Sui, Y., Guo, Y., Li, B., Meng, L., & Ma, Y. (2022). Estimation of sugar content in sugar beet root based on uav multi-sensor data. *Computers and Electronics in Agriculture*, *203*, 107433. https://doi.org/10.1016/j.compag.2022.107433

Wei, X., He, J.-C., Ye, D.-P., & Jie, D.-F. (2017). Navel orange maturity classification by multispectral indexes based on hyperspectral diffuse transmittance imaging. *Journal of Food Quality*, *2017*. https://doi.org/10.1155/2017/1023498

Yang, Y., Ren, J., Zheng, X.-Q., Zhao, L.-Y., & Li, M.-M. (2014). Rapid determination of beet sugar content using near infrared spectroscopy. *Guang pu xue yu guang pu fen xi= Guang pu*, *34*(10), 2728-2731.

Yu, Y., Zhang, Q., Huang, J., Zhu, J., & Liu, J. (2021). Nondestructive determination of ssc in korla fragrant pear using a portable near-infrared spectroscopy system. *Infrared Physics & Technology*, *116*, 103785. https://doi.org/10.1016/j.infrared.2021.103785

Zhang, C., Liu, F., Kong, W., & He, Y. (2015). Application of visible and near-infrared hyperspectral imaging to determine soluble protein content in oilseed rape leaves. *Sensors*, *15*(7), 16576-16588. https://doi.org/10.3390/s150716576

Zhang, J., Guo, Z., Ren, Z., Wang, S., Yin, X., Zhang, D., Wang, C., Zheng, H., Du, J., & Ma, C. (2023). A non-destructive determination of protein content in potato flour noodles using near-infrared hyperspectral imaging technology. *Infrared Physics & Technology*, *130*, 104595. https://doi.org/10.1016/j.infrared.2023.104595

Zhang, X., Sun, J., Li, P., Zeng, F., & Wang, H. (2021). Hyperspectral detection of salted sea cucumber adulteration using different spectral preprocessing techniques and svm method. *LWT*, *152*, 112295. https://doi.org/10.1016/j.lwt.2021.112295